

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号
特開2000-123033
(P2000-123033A)

(43)公開日 平成12年4月28日(2000.4.28)

(51)Int.Cl.⁷
G 0 6 F 17/30

識別記号

F I
G 0 6 F 15/403

テーマコード(参考)

3 5 0 C 5 B 0 7 5
3 4 0 A

審査請求 未請求 請求項の数13 O L (全 17 頁)

(21)出願番号 特願平10-292151

(22)出願日 平成10年10月14日(1998.10.14)

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 梅木 秀雄

神奈川県川崎市幸区柳町70番地 株式会社
東芝柳町工場内

(72)発明者 下郡 信宏

神奈川県川崎市幸区柳町70番地 株式会社
東芝柳町工場内

(74)代理人 100058479

弁理士 鈴江 武彦 (外6名)

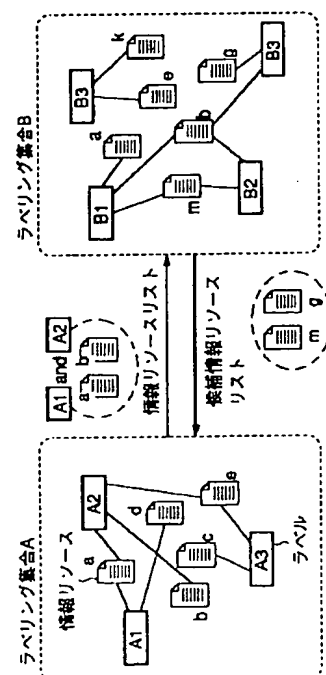
最終頁に続く

(54)【発明の名称】 情報検索方法および情報検索装置および記録媒体

(57)【要約】

【課題】ラベル付けにより整理分類した結果を用いて、関心のある情報と関連する未知の情報リソースを効率良く検索および収集できる情報検索方法およびそれを用いた情報検索装置を提供する。

【解決手段】情報リソースと該情報リソースに付与したラベルとの対応関係にて特徴づけられる複数の情報リソースからなる複数のラベリング集合のうちの1つ(ラベリング集合A)から選択された該ラベルの論理式にて特徴付けられる1または複数の検索用情報リソース(a、b)を、他のラベリング集合Bから検索し、その検索された検索用情報リソースに付されたラベル(B1、B2、B4)について第1の評価値を求め、該第1の評価値に基づいて、該検索用情報リソース以外の情報リソース(m、g)の第2の評価値を求め、該第2の評価値に基づき選択した該検索用情報リソースに関連する情報リソースを呈示する。



【特許請求の範囲】

【請求項 1】 情報リソースと該情報リソースに付与したラベルとの対応関係にて特徴づけられる複数の情報リソースからなる複数のラベリング集合のうちの 1 つから選択された該ラベルの論理式にて特徴付けられる 1 または複数の検索用情報リソースを、他の 1 または複数のラベリング集合のそれぞれから検索し、

その検索された検索用情報リソースに付されたラベルについての該検索用情報リソースと該ラベルとの結びつきを評価する第 1 の評価値を求め、

該第 1 の評価値に基づいて、該検索用情報リソース以外の情報リソースの該検索用情報リソースとの関連の度合いを評価する第 2 の評価値を求め、

該第 2 の評価値に基づき選択した該検索用情報リソースに関連する情報リソースを呈示することを特徴とする情報検索方法。

【請求項 2】 複数のラベリング集合のそれぞれから求められた情報リソースの第 2 の評価値を該複数のラベリング集合のそれぞれに予め定められた重み係数に基づき統合して、その統合結果に基づき前記検索用情報リソースに関連する情報リソースを選択することを特徴とする請求項 1 記載の情報検索方法。

【請求項 3】 前記検索用情報リソースに付されたラベルについて、そのラベルの付された該検索用情報リソース 1 つにつき所定値を累積加算することにより、該ラベルの第 1 の評価値を算出することを特徴とする請求項 1 記載の情報検索方法。

【請求項 4】 前記検索用情報リソースに付されたラベルについて、そのラベルの付された該検索用情報リソースの重要度に基づき該ラベルの第 1 の評価値を算出することを特徴とする請求項 1 記載の情報検索方法。

【請求項 5】 前記検索用情報リソース以外の情報リソースについて、その情報リソースに付されているラベルの前記第 1 の評価値を累積加算することにより、該情報リソースの第 2 の評価値を算出することを特徴とする請求項 1 記載の情報検索方法。

【請求項 6】 前記検索用情報リソース以外の情報リソースについて、その情報リソースに付されているラベルの前記第 1 の評価値と該ラベルの重要度とに基づき、該情報リソースの第 2 の評価値を算出することを特徴とする請求項 1 記載の情報検索方法。

【請求項 7】 情報リソースと該情報リソースに付与したラベルとの対応関係にて特徴づけられる複数の情報リソースからなる複数のラベリング集合のうちの 1 つから選択された該ラベルの論理式にて特徴付けられる 1 または複数の検索用情報リソースを、他の 1 または複数のラベリング集合のそれぞれから検索する検索手段と、

この検索手段で検索された該検索用情報リソースに付されたラベルについての該検索用情報リソースと該ラベルとの結びつきを評価する第 1 の評価値を求めるラベル評

価手段と、

前記第 1 の評価値に基づいて、該検索用情報リソース以外の情報リソースの該検索用情報リソースとの関連の度合いを評価する第 2 の評価値を求める情報リソース評価手段と、

前記第 2 の評価値に基づき選択した該検索用情報リソースに関連する情報リソースを呈示する呈示手段と、を具備したことを特徴とする情報検索装置。

【請求項 8】 複数のラベリング集合のそれぞれから求められた情報リソースの第 2 の評価値を該複数のラベリング集合のそれぞれに予め定められた重み係数に基づき統合して、その統合結果に基づき前記検索用情報リソースに関連する情報リソースを選択することを特徴とする請求項 7 記載の情報検索装置。

【請求項 9】 前記検索用情報リソースに付されたラベルについて、そのラベルの付された該検索用情報リソース 1 つにつき所定値を累積加算することにより該ラベルの第 1 の評価値を算出することを特徴とする請求項 7 記載の情報検索装置。

【請求項 10】 前記検索用情報リソースに付されたラベルについて、そのラベルの付された該検索用情報リソースの重要度に基づき該ラベルの第 1 の評価値を算出することを特徴とする請求項 7 記載の情報検索装置。

【請求項 11】 前記検索用情報リソース以外の情報リソースについて、その情報リソースに付されているラベルの前記第 1 の評価値を累積加算することにより、該情報リソースの第 2 の評価値を算出することを特徴とする請求項 7 記載の情報検索装置。

【請求項 12】 前記検索用情報リソース以外の情報リソースについて、その情報リソースに付されているラベルの前記第 1 の評価値と該ラベルの重要度とに基づき、該情報リソースの第 2 の評価値を算出することを特徴とする請求項 7 記載の情報検索装置。

【請求項 13】 情報リソースと該情報リソースに付与したラベルとの対応関係にて特徴づけられる複数の情報リソースからなる複数のラベリング集合のうちの 1 つから選択された該ラベルの論理式にて特徴付けられる 1 または複数の検索用情報リソースを、他の 1 または複数のラベリング集合のそれぞれから検索させる検索手段と、この検索手段で検索された該検索用情報リソースに付されたラベルについての該検索用情報リソースと該ラベルとの結びつきを評価する第 1 の評価値を算出させるラベル評価手段と、

前記第 1 の評価値に基づいて、前記検索用情報リソース以外の情報リソースの該検索用情報リソースとの関連の度合いを評価する第 2 の評価値を算出させる情報リソース評価手段と、

前記第 2 の評価値に基づき選択した前記検索用情報リソースに関連する情報リソースを呈示させる呈示手段と、を実行するプログラムを記録した機械読み取り可能な記

録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、複数のユーザが参照可能な文書や画像などの情報リソースに対して、関心のある情報リソース集合と関連した未知の情報リソースを検索、収集する情報収集方法およびそれを用いた情報検索装置に関し、例えば、映画や音楽などのコンテンツに対して、嗜好の似た人から関連した未知の情報を取得する協調型フィルタリング技術、WWW(World Wide Web)ブラウザのブックマーク機能の公開・共有化によって、ネットワーク上におけるグループ(コミュニティ)内での情報共有を促進するための技術などに関わる。

【0002】

【従来の技術】ラベル付けを利用した個人の情報分類あるいは整理方式は、従来からあるメモ、アノテーション(注釈)機能と類似した用途としてすでに存在している。これらの手法の特徴としては、ユーザがつけたラベルの名前やラベルにつけられた属性値を利用して情報を分類、検索することである。このとき、利用範囲が個人に限られている場合にはあまり問題にならないが、ユーザが勝手につけたラベル名あるいはメモ内容を直接利用して他人との情報共有や検索を行うことを考えると、ラベル名やメモで利用した言葉の統一性や整合性が問題になる。

【0003】これを回避するため、システム構築の際には、ラベル名や用語を統一して名前空間を制限するか、シソーラスや単語データベースにより、名前の関連付けをシステム側が処理するかのいずれかを選択することが多い。

【0004】後者の場合には、規模の大きな辞書を用意して検索の精度を上げてみてもなお、単語の多義性や曖昧性の問題は残る。また、ドキュメントの文章を解析してキーワードを抽出し、情報検索に用いることも、Webのドキュメントを検索する際にしばしば利用される。

【0005】しかし、文章中の単語を抽出するだけでは、そこに単語が存在していることはわかるが、そのドキュメントがどのような観点で書かれていて、有用かどうかなどといった分析は困難である。まして、画像や音声、動画情報のみの情報リソースの場合にはキーワード解析はできない。いずれにしても、そのドキュメントがユーザにとってどれくらい有用なのか、有用だとすればどのような観点から有用なのかといった評価や分類は各ユーザごとに異なるものである。

【0006】ユーザが行った評価情報を積極的に利用するシステムもある。その代表的な例は、推薦エンジン(Recommendation Engine)と呼ばれる協調型フィルタリング(Collaborative Filtering)を利用したシステムであ

る。これはインターネット上でビデオや音楽、書籍を宣伝、販売する際に、ユーザの好みにあったリストを精度よく呈示する手法の1つである。具体的には、たとえば多くの映画に対して大勢のユーザが与えた数値的な評価(好きか嫌いか、あるいはどのくらい好きか)をデータベース化し、各人の嗜好性の相関を計算することにより、そのユーザの好みと類似したグループから、そのユーザが好みそうな映画のタイトルのリストを呈示するといったものである。推薦エンジンの代表的な例としては、GroupLens推薦エンジン(USA、Net Perception社)が挙げられる。現在の推薦エンジンは、評価用の定型ラベル(たとえば「好き」から「嫌い」までを数値化したマークシート的な評価用ラベル)を映画や音楽などのデータに付与していると思えることができる。しかし、コンテンツに対する評価としては、好きか嫌いかなどの一面的な捉え方しかできないのが現実である。

【0007】

【発明が解決しようとする課題】このように、WWW上の「Yahoo」や「Infoseek」などの通常のキーワードに基づく検索エンジンでは、検索すること、情報を集めて整理するという行為(ブックマークやファイリング)とが完全に分離しているため、ある事柄に関して情報を集めたとしても、それを検索の精度を向上させるために利用することができないという問題点があった。

【0008】そこで、本発明は、ネットワーク上の複数のユーザが参照可能な文書や画像といった任意の情報リソースに対して、各ユーザが自由な観点に基づいて収集し、分類した結果得られる情報を使って、関心のある情報と関連する未知の情報リソースを効率良く検索および収集できる情報検索方法およびそれを用いた情報検索装置を提供することを目的とする。

【0009】

【課題を解決するための手段】(1)本発明の情報検索方法(請求項1)は、情報リソースと該情報リソースに付与したラベルとの対応関係にて特徴づけられる複数の情報リソースからなる複数のラベリング集合のうちの1つ(例えば、ラベリング集合A)から選択された該ラベルの論理式にて特徴付けられる1または複数の検索用情報リソース(コレクタに保存される)を、他の1または複数のラベリング集合(例えばラベリング集合B、C)のそれぞれから検索し、その検索された該検索用情報リソースに付されたラベルについての該検索用情報リソースと該ラベルとの結びつきを評価する第1の評価値(ラベルスコア)を求め、該第1の評価値に基づいて、前記検索用情報リソース以外の情報リソースの該検索用情報リソースとの関連の度合いを評価する第2の評価値(情報リソーススコア)を求め、該第2の評価値に基づき選択した前記検索用情報リソースに関連する情報リソース

を呈示することを特徴とする。

【0010】本発明によれば、各ユーザが自由な観点に基づいて収集した情報リソースにラベリング（ラベル付け）により整理分類した結果を用いて、関心のある情報と関連する未知の情報リソースを効率良く検索および収集できる。

【0011】（請求項2）好ましくは、複数のラベリング集合のそれぞれから求められた情報リソースの第2の評価値を該複数のラベリング集合のそれぞれに予め定められた重み係数に基づき統合して、その統合結果に基づき前記検索用情報リソースに関連する情報リソースを選択する。

【0012】（請求項3）好ましくは、前記検索用情報リソースに付されたラベルについて、そのラベルの付された該検索用情報リソース1つにつき所定値を累積加算することにより、該ラベルの第1の評価値を算出する。

【0013】（請求項4）好ましくは、前記検索用情報リソースに付されたラベルについて、そのラベルの付された該検索用情報リソースの重要度に基づき該ラベルの第1の評価値を算出する。

【0014】（請求項5）好ましくは、前記検索用情報リソース以外の情報リソースについて、その情報リソースに付されているラベルの前記第1の評価値を累積加算することにより、該情報リソースの第2の評価値を算出する。

【0015】（請求項6）好ましくは、前記検索用情報リソース以外の情報リソースについて、その情報リソースに付されているラベルの前記第1の評価値と該ラベルの重要度とに基づき該情報リソースの第2の評価値を算出する。

【0016】（2）本発明の情報検索装置（請求項7）は、情報リソースと該情報リソースに付与したラベルとの対応関係にて特徴づけられる複数の情報リソースからなる複数のラベリング集合のうちの1つから選択された該ラベルの論理式にて特徴付けられる1または複数の検索用情報リソースを、他の1または複数のラベリング集合のそれぞれから検索する検索手段と、この検索手段で検索された該検索用情報リソースに付されたラベルについての該検索用情報リソースと該ラベルとの結びつきを評価する第1の評価値を求めるラベル評価手段と、前記第1の評価値に基づいて、前記検索用情報リソース以外の情報リソースの該検索用情報リソースとの関連の度合いを評価する第2の評価値を求める情報リソース評価手段と、前記第2の評価値に基づき選択した前記検索用情報リソースに関連する情報リソースを呈示する呈示手段と、を具備したことを特徴とする。

【0017】本発明によれば、各ユーザが自由な観点に基づいて収集した情報リソースにラベリング（ラベル付け）により整理分類した結果を用いて、関心のある情報と関連する未知の情報リソースを効率良く検索および収

集できる。

【0018】（請求項8）好ましくは、複数のラベリング集合のそれぞれから求められた情報リソースの第2の評価値を該複数のラベリング集合のそれぞれに予め定められた重み係数に基づき統合して、その統合結果に基づき前記検索用情報リソースに関連する情報リソースを選択する。

【0019】（請求項9）好ましくは、前記検索用情報リソースに付されたラベルについて、そのラベルの付された該検索用情報リソース1つにつき所定値を累積加算することにより、該ラベルの第1の評価値を算出する。

【0020】（請求項10）好ましくは、前記検索用情報リソースに付されたラベルについて、そのラベルの付された該検索用情報リソースの重要度に基づき該ラベルの第1の評価値を算出する。

【0021】（請求項11）好ましくは、前記検索用情報リソース以外の情報リソースについて、その情報リソースに付されているラベルの前記第1の評価値を累積加算することにより、該情報リソースの第2の評価値を算出する。

【0022】（請求項12）好ましくは、前記検索用情報リソース以外の情報リソースについて、その情報リソースに付されているラベルの前記第1の評価値と該ラベルの重要度とに基づき、該情報リソースの第2の評価値を算出する。

【0023】（3）本発明の記録媒体（請求項13）は、情報リソースと該情報リソースに付与したラベルとの対応関係にて特徴づけられる複数の情報リソースからなる複数のラベリング集合のうちの1つから選択された該ラベルの論理式にて特徴付けられる1または複数の検索用情報リソースを、他の1または複数のラベリング集合のそれぞれから検索させる検索手段と、この検索手段で検索された該検索用情報リソースに付されたラベルについての該検索用情報リソースと該ラベルとの結びつきを評価する第1の評価値を算出させるラベル評価手段と、前記第1の評価値に基づいて、該検索用情報リソース以外の情報リソースの該検索用情報リソースとの関連の度合いを評価する第2の評価値を算出させる情報リソース評価手段と、前記第2の評価値に基づき選択した該検索用情報リソースに関連する情報リソースを呈示させる呈示手段と、を実行するプログラムを記録した機械読み取り可能な記録媒体であり、本発明によれば、上記手段を、例えば、ラベルに対する操作を行うために画面

（パレット）、情報検索のために利用するコレクタの作成や検索結果の呈示を行うためのファインダ等のユーザインタフェース画面を呈示するためのディスプレイ装置と、ハードディスク装置等の記憶装置を具備した（好ましくは、さらに、インターネット等のネットワークを介して通信可能な）パーソナルコンピュータ等のコンピュータに実行させることのできるプログラムとして、磁気

ディスク（フロッピーディスク、ハードディスクなど）、光ディスク（CD-ROM、DVDなど）、半導体メモリなどの記録媒体に格納して頒布することにより、各ユーザが自由な観点に基づいて収集した情報リソースにラベリング（ラベル付け）により整理分類した結果を用いて、関心のある情報と関連する未知の情報リソースを効率良く検索および収集できる。

【0024】

【発明の実施の形態】以下、本発明の実施の形態について図面を参照して説明する。ここで、以下の説明で用いる用語の定義を行う。情報リソースとは、例えば、例えば、音声、テキスト、静止画、動画像、HTML文書等の情報ファイルである。

【0025】複数の情報リソースに対して任意の基準や観点に基づいた主観的な自由な名前をもつラベルを付与する行為を「ラベリング」と呼び、ラベリングの結果得られるラベルと情報リソースの対応関係の集合を「ラベリング集合」と呼ぶことにする。

（第1の実施形態）図1は、本発明の第1の実施形態に係る情報検索方法を用いた情報検索装置の基本的な構成を概略的に示したもので、大きく分けて、情報検索処理部1と、データ格納部50とから構成され、情報検索処理部1は、さらに、検索用情報リソースリスト取得部10と、ラベリング集合選択部20と、情報検索部30と、候補情報リソース呈示部40とから構成されている。

【0026】ここでは、例えば、ブラウザを用いてインターネット上から収集した、あるいは自分で作成した情報リソースに対し、ユーザが自分が整理したいあるいは記録しておきたい基準や観点に基づいて自由に決めたラベルを付して、既にラベリングされており、そのラベルに関する情報、ラベルを付けた情報リソースに関する情報等は、データ格納部50に格納されているものとする。

【0027】図2は、2つのラベリング集合A、Bを概念的に示したものである。ここでは、例えば、ラベリング集合Aとラベリング集合Bは、それぞれ異なるユーザがもつラベリング集合を示しているが、1人のユーザが複数のラベリング集合A、Bをもつ場合であってもよい。

【0028】1つのラベリング集合には、その内部で定義されたラベルの集合と、そのラベルが付けられた任意の情報リソースの集合が存在する。たとえば、ラベリング集合A内には、ラベル集合{A1、A2、A3}とそれらのラベルが付けられた情報リソースの集合{a、b、c、d、e}がある。

【0029】図2では、ラベルと情報リソース間の対応づけ、つまりある情報リソースに対してあるラベルが付けられているという関係はラベルと情報リソースを結ぶ線で表現されている。

【0030】ラベリング集合Bにおいてもラベリング集合Aと同様に、ラベル集合{B1、B2、B3、B4}が存在し、情報リソース{a、b、e、f、g、k}に対してこれらのラベルが付けられている。名前が同じ情報リソースは同じ実体を参照していることを示している。このとき、たとえばラベリング集合Aにおいて、ラベルA1とラベルA2がともに付いている情報リソースの集合は、ラベルA1とラベルA2の両方に関する事柄を含んでいると考えてよい。

【0031】このように、ラベリングの結果、情報リソースはラベルの論理式による組み合わせによって、様々な切口で分類されている。ラベルの論理式とは、例えば、「A1 and A2」という形で表現したものであり、R(A1 and A2)とは、ラベルA1とラベルA2とがともに付いた情報リソースの集合を意味する。例えば、図2に示したラベリング集合A、Bでは、 $R(A1 \text{ and } A2) = \{a, b\}$ である。

【0032】本発明の主な目的は、このようなラベルの論理式によって特徴づけられる情報リソースの集合を検索キーとして、別のラベリング集合からそれらの情報リソースに関連する未知の情報リソースを検索し、収集することである。

【0033】図1の検索用情報リソースリスト取得部10では、ラベリング集合内のラベルの論理式によって表現された情報リソースのリストを取得し、これを検索のキーとして利用する。この処理は、たとえば図2のラベリング集合Aにおいて、前述したように、ラベルの論理式「A1 and A2」を指定して情報リソースリスト{a、b}を得ることに相当する。基本的には、情報リソースリスト{a、b}はラベルの論理式によって得るが、必要に応じて、この情報リソースリストから特定の情報リソースを取り除いたり、付け加えることもできる。

【0034】検索のための情報リソースの集合を表した情報リソースリスト（具体的には、情報リソースの識別情報（ID）のリストで、メインリストともいう）を取得したあとで、つぎにこれらの情報リソースに関連した情報を検索する対象（他のラベリング集合）を決める。そのため、ラベリング集合選択部20では、検索の対象となる他のラベリング集合を選択し、検索用情報リソースリスト取得部10で得られた情報リソースリストを相手のラベリング集合に通知する。すなわち、検索対象として、社内であれば、特定の部署を指定して、当該部署に所属する人の有するラベリング集合に限定したり、特定の人物を指定して、当該人物の有するラベリング集合に限定したりといった検索範囲を限定する。なお、その必要がなければ、必ずしも検索範囲を限定する必要はない。

【0035】図2の例では、ラベリング集合Bが選択され、情報リソースリスト{a、b}がラベリング集合B

に通知されることに相当する。一般に選択すべきラベリング集合は複数であるが、暗黙的に検索の発信源であるラベリング集合（ここでは、ラベリング集合A）以外のすべてのラベリング集合が検索対象となる。

【0036】次に、情報検索部30において各ラベリング集合ごとに関連情報を検索する。情報検索部30は、ラベルスコア算出部31と、情報リソーススコア算出部32と、候補情報リソースリスト作成部33とから構成される。

【0037】ここで、ラベルスコアとは、検索用に与えた情報リソースの集合Rsに対して、あるラベリング集合内で定義されたラベルとの結び付きの度合いを表す。また、情報リソーススコアは、ラベルスコアに基づいて計算される、Rsとそれ以外の情報リソースとの関連性を計るための指標である。簡単に説明すれば、与えた情報リソースの集合Rsと同じラベル付けがされている情報リソースほどスコアが高いということになり、ラベル付けという分類上、もとのRsとの関連性が高い情報リソースであるといえる。

【0038】図3は、ラベルスコア算出部31と情報リソーススコア算出部32の処理動作を示したフローチャートで、図4は、図3のステップS2からステップS4までの処理（ラベルスコアの算出処理）を説明するための概念図で、図5は、図3のステップS5からステップS7までの処理（情報リソーススコアの算出処理）を説明するための概念図である。

【0039】まず、ラベルスコア算出部31の処理動作（ラベルスコアの算出処理）について、図3と図4を参照して説明する。図4のステップS1で、ラベルスコアと情報リソーススコアを、例えば、「0」という値に初期化しておく。つぎに、ステップS2で、検索用の情報リソース集合をRsとしたとき、Rsのすべての要素（情報リソース）を調べたかどうかをチェックし、まだ調べるべき要素が残っている場合には、ステップS3に進み、Rsから1つの要素を取り出して、それをrとする。ステップS4では、当該取り出した情報リソースrに付けられた全ラベルのラベルスコアに所定の数値（例えば、「1」）を加算する。

【0040】図4において、Rsとして情報リソース{a, b}が与えられたとき、ラベリング集合B内において、Rsの各要素（情報リソース）に付けられたラベルに対して、そのラベルが付けられている1つの情報リソースにつき、全て同等に数値「1」がラベルスコアに加算されている。ラベルB1は、Rsとして与えられた情報リソースa, bのラベルであり、ラベルB2は、Rsとして与えられた情報リソースbのラベルであり、ラベルB4は、Rsとして与えられた情報リソースbのラベルであるので、ラベルB1、B2、B4のラベルスコアSc(B1)、Sc(B2)、Sc(B4)は、それぞれ、Sc(B1) = 1 + 1 = 2、Sc(B2) = 1、S

c(B4) = 1 となる。

【0041】ここでは、どのラベルに対しても情報リソース1つにつき加算する数値は、全て同等の「1」にしているが、この理由は、ラベリング集合Aから選んだ情報リソース集合Rsの要素（情報リソース）間に重要度などの順位付けが存在しない場合には、ラベリング集合B上でRsの要素に付けられたラベルに対して一律に同じ数値を加算することは妥当であると考えられるからである。

【0042】もし、ラベリング集合A上でラベルA1とラベルA2に重要度の違いがあるならば、例えば、ラベリング集合Aから検索用の情報リソースの集合として、 $Rs = R(A1 \text{ or } A2) = \{a, b, d, e\}$ （「or」は「または」を意味する論理演算記号で、この場合、ラベルA1かラベルA2のうちの少なくとも一方が付されている情報リソースの集合を意味する）を選んだとき、付与されたラベルの重要度に応じて、この4つの情報リソースに重要度を割り当てることができる。したがって、ラベリング集合B上でラベルスコアを計算するときに、各情報リソースに、この重要度に応じて各情報リソースに高低をつけた数値を割り当てて、その情報リソースに付されたラベルに当該値を加算するようにしてもよい。

【0043】次に、情報リソーススコア算出部32の処理動作（情報リソーススコアの算出処理）について、図3と図5を参照して説明する。情報リソーススコア算出部32では、ラベルスコア算出部31で得られた各ラベルのラベルスコアの値をそのラベルが付けられたRs以外の情報リソースについて加算したものを情報リソーススコアとする。ここで、ラベリング集合B上で検索用の情報リソースの集合Rsに付けられたラベルの集合をL(Rs)とする。

【0044】図3のステップS5では、L(Rs)の要素（ラベル）のすべてを調べたかどうかをチェックし、まだ調べるべきラベルが残っていれば、ステップS6に進み、L(Rs)から1つの要素を取り出して、それを入とする。ステップS7では、当該取り出したラベル入が付けられたRs以外の情報リソースのスコアに入のラベルスコアを加算する。

【0045】図5において、ラベルスコア算出部31で得られた各ラベルB1、B2、B4のラベルスコアの値をそのラベルが付けられたRs以外の情報リソース

{m, g}について加算する。たとえば、情報リソースmには、ラベルB1とラベルB2とが付いている。情報リソースmのスコアSc(m)は、それに付けられているラベルB1とラベルB2のラベルスコアSc(B1)、Sc(B2)を加算したものになっている。つまり、 $Sc(m) = Sc(B1) + Sc(B2) = 2 + 1 = 3$ となる。情報リソースgのスコアに関しても同様に、ラベルB4のラベルスコアから、 $Sc(g) = Sc$

(B4) = 1となる。

【0046】このとき、ラベルに重要度が与えられていたとすれば、情報リソーススコアを求める際に、各ラベルスコアからの寄与を重要度によって変えることもできる。例えば、各ラベルに、その重要度に応じて高低をつけた数値（例えば、重要度の高いものには「2」、重要度の低いものには「1」）を割り当てて、その数値と情報リソースのスコアとを乗じてから、加算するようにしてもよい。例えば、ラベルB1は重要度が高いので「2」を割り当て、ラベルB2は重要度が低いので「1」を割り当てるとすると、情報リソースmのスコアは、 $S_c(m) = \{S_c(B1) \cdot 2\} + \{S_c(B2) \cdot 1\} = 4 + 1 = 5$ となる。

【0047】候補情報リソースリスト作成部33では、情報リソーススコアの値が高い情報リソースを候補情報リソースリスト（候補リスト）として選び、そのスコアの値に基づいて情報リソースを順位付けしたリストを作成する。例えば、図5の例では、情報リソーススコアが「0」でない情報リソースは、mとgだけなのでこれらをラベリング集合Bから検索された候補情報リソースの集合 $R_c(B) = \{m, g\}$ とし、各々の情報リソーススコアの値によって、1番目の候補情報リソースをm、2番目の候補情報リソースをgとする。

【0048】最後に、候補情報リソース呈示部40で、候補情報リソースリスト作成部33で作成した順位付けされた候補情報リソースのリストを検索を発した元のラベリング集合に対して呈示する。

【0049】以上は、1つのラベリング集合を検索の対象として想定し、説明してきたが、複数のラベリング集合にわたる情報検索の場合は、基本的には個々の検索対象のラベリング集合で得られた情報検索の結果を重ね合わせる（統合する）ことになる。

【0050】図6は、複数（例えば4つ）のラベリング集合A～Dから情報リソースを検索する場合を概念的に示したものである。例えば、上記同様、ラベリング集合Aから検索が発せられると、ラベリング集合B～D毎に上記同様にして候補情報リソースの集合 $R_c(B)$ 、 $R_c(C)$ 、 $R_c(D)$ が得られたとする。このとき、これら候補情報リソースの集合の論理和をとり、同じ候補情報リソースに対する情報リソーススコアについては、各ラベリング集合ごとに予め与えられた重み係数を掛けたものを足し合わせて、統合スコアを求める。

【0051】具体的には、たとえばラベリング集合B、C、Dのそれぞれから候補情報リソースとして、それぞれ、 $R_c(B) = \{m(3), g(1)\}$ 、 $R_c(C) = \{g(2), m(1), p(1)\}$ 、 $R_c(D) = \{m(2), p(2), e(1)\}$ が得られたとする。ただし、各要素（情報リソース）の括弧内の数値は情報リソーススコアの値を表すとする。このとき、これら候補情報リソースの論理和をとり、最終的な候補情報リソ

ースの集合は、 $R_c = \{m, g, p, e\}$ である。ラベリング集合B、C、Dの重み係数として、 K_B 、 K_C 、 K_D が与えられていたとすると、候補情報リソースmの統合スコアは、各ラベリング集合B、C、Dにおける情報リソーススコアに各ラベリング集合の重み係数を乗じてから加算すればよいので、

$$S_c(m) = 3 \cdot K_B + 1 \cdot K_C + 2 \cdot K_D$$

となる。通常は、各ラベリング集合の重み係数はすべて同じと考えて良いので、その場合には、各候補情報リソースの統合スコアは、各ラベリング集合において計算された情報リソーススコアの単純な和でよいことになる。

（第2の実施形態）図7は、本発明の情報検索方法を用いて情報を検索および収集する場合を概念的に示したものである。ユーザ80は、情報リソースに対して自分で定義したラベルを付けていく。ある情報リソースの集合に関連した情報を検索したい場合には、まず、その情報リソース集合を特徴づけるラベルの論理式を検索キーとして後述するコレクタ81と呼ばれる仮想的なフォルダを作成することにより、自分のラベリング集合から対応する情報リソースを抽出する。つぎに、それらの情報リソースリストを検索のキーとして、他の複数のラベリング集合83に対して検索を行う。検索結果84は候補情報リソースの順位付けされたリストとして得られる。必要に応じて候補情報リソースの中から幾つかをコレクタに追加することが可能である。

【0052】図8は、本発明の第2の実施形態に係る情報検索装置を用いて、例えば、ネットワークを介して情報検索を行う場合を概念的に示したものである。情報リソースは、イントラネット及びインターネット上にあって、複数のユーザ（例えば2人のユーザ α 、 β ）がアクセス可能であり、図8では情報リソースは、データ格納部402、422、432に格納されている。この例ではラベリングを行う主体はユーザすなわち人であることを想定している。

【0053】例えば、ユーザ α は、クライアントとしての情報検索装置（以下、簡単に端末と呼ぶこともある）421を利用して、自分がアクセスした情報に対してラベリングを行う。ユーザに関する情報、ユーザが作成したラベルに関する情報、ユーザがラベルを付けた情報リソースに関する情報、情報リソースを格納したコレクタに関する情報等は、各クライアント（端末421、431）のデータ格納部422、432に格納されるとともに、ネットワーク400を介して、ラベリングの情報を管理、処理するラベリングサーバ411に送られ、ラベリングDB412にも保存される。クライアントからラベリング集合間の情報検索を行う場合には、ラベリングサーバ411に対して必要な処理を要求する。なお、ラベリングサーバ411で保持・管理されている情報は、各クライアント側で保持しているデータと整合性を保つように管理されている。各ユーザの端末421、431

では、各ユーザがそれぞれ独自にラベリングした結果としてのラベルに関する情報、ユーザがラベルを付けた情報リソースに関する情報、情報リソースを格納したコレクタに関する情報等を保持するとともに、各端末421、431にあるラベリングインタフェースを介して、ラベル付けや自分や他人のラベリングに対する情報検索を行うようになっている。ラベリングインタフェースは、パレット、ファインダ、コレクタという3つの基本コンポーネントから構成されている。

【0054】図9は、本発明の第2の実施形態に係る(図8の端末421、431に相当する)情報検索装置の要部の構成を概略的に示したもので、前述の第1の実施形態に係る情報検索装置の構成をより具体的に示したものである。

【0055】図9に示す情報検索装置は、大きく分けて、ユーザインタフェース部501と処理部502と、データ格納部506と、送受信部507とから構成されている。

【0056】処理部502は、ラベリング処理を行うためのラベリング処理部503と、コレクタ作成処理を行うためのコレクタ作成処理部504と、情報検索処理を行うための情報検索処理部505とから構成されている。

【0057】ラベリング処理部503は、ラベル定義部510とラベル付与部520とから構成され、コレクタ作成処理部504は、コレクタ検索式設定部530とメインリスト取得部540とメインリスト操作部550とから構成され、情報検索処理部505は、コレクタ選択部560と情報検索部570と候補情報リソースリスト呈示部580とから構成されている。

【0058】ユーザインタフェース部501は、上記各部でラベリング処理、コレクタ作成処理、情報検索処理を実行する際に必要な情報の呈示および入力を行うためのもので、例えば、ディスプレイ装置や、マウス、キーボード等の入力装置等から構成されている。

【0059】データ格納部506には、ユーザにより収集された情報リソースと、当該情報リソースに対しラベリングした結果であるラベル情報等を格納する。また、コレクタ作成処理部504で作成されたコレクタも格納される。

【0060】送受信部507は、コレクタ作成処理部504で作成されたコレクタを検索要求としてネットワーク400を介してラベリングサーバ411へ送信したり、検索結果をネットワーク400を介して受信したりする。

【0061】図10は、図9の情報検索装置の動作を、ユーザインタフェース部501にて呈示されるユーザインタフェース画面(ラベルの操作を行うパレット610と、ラベルやコレクタごとに関係した情報リソースのリストを表示するファインダ620)と、主に情報検索の

ために利用するコレクタ630との関連に基づき概念的に示したものである。

【0062】図10に示すように、パレット610では、ラベルの操作を行い、ファインダ620では、ラベルやコレクタごとに関係した情報リソースのリストを表示し、情報検索のために用いるコレクタ630を作成し、検索結果の呈示を行う。

【0063】まず、図9のラベリング処理部503について説明する。ラベリング、すなわち、ある情報リソースに対してユーザが定義したラベルを付与するという処理は、ラベル定義部510とラベル付与部520にて実行される。これらの処理は、パレット610のようなインタフェース画面をユーザインタフェース部501に呈示して行う。

【0064】ラベル定義部510では、ユーザが自由にラベルの名前を決定し、ラベルを新規に作成する。また必要に応じて、ラベル名を変更したり、ラベルの消去を行う。ラベル付与部520では、情報リソースに対してラベルを付与する。

【0065】パレット610は、別のブラウザ600と呼ばれるインタフェース画面を介して、現在表示している情報リソースのID(例えば、URL)を取得し、その取得した情報リソースに、マウス操作などによって付与すべきラベルを選択してラベル付けを行うためのインタフェース画面である。

【0066】図11は、ユーザインタフェース部501にて呈示されるパレット610の基本画面の一例を示したものである。図11に示したように、パレットの基本画面には、ユーザが作成したラベル(「料理」「健康」「役場」「写真」「香港映画」「論文」)が配置されている。

【0067】新たにラベルを作成しようとするときは、マウス等を用いて、「New」というボタン614を押下すれば、任意の位置にラベルの入力領域が呈示されるので、そこに所望のラベル名を入力すればよい。また、既に作成されているラベルを消去しようとするときは、消去しようとするラベルを選択して(選択されたラベルは、例えば、反転表示する)、「Delete」というボタン615を押下すればよい。

【0068】ユーザは、ラベルをドラッグ&ドロップ操作によって任意の位置に動かすことができるため、例えば、ジャンルの類似しているラベル(例えば「料理」と「健康」)を集めてグループ化することができる。

【0069】パレット610上で情報リソースにラベルを貼り付けたい場合は、ユーザは、例えば、情報リソースID指定領域611に、所望の情報リソースID(例えば、URL)を入力し、あるいは、下向きの矢印が表記されたボタン616を押下することにより表示された選択可能な情報リソースのIDのリストから所望の情報リソースIDを選択する。あるいは、ブラウザ600上

に開かれている情報リソースのIDが自動的に情報リソースID指定領域611に取り込まれるようになっていてもよい。次に、パレット610上に呈示されているラベルのうち所望のものを選択して（選択されたラベルは、画面上で反転表示される）、「Labeling」というラベルの貼り付け指示用のボタン612を押下すればよい。

【0070】一度貼り付けたラベルを変更したい場合は、情報リソースID指定領域611にラベルを変更したい情報リソースのIDを入力（あるいは選択）すれば、その情報リソースに既に貼り付けてあるラベルが画面上で選択状態（例えば画面上で反転表示する）になるので、選択状態を変更して再び「Labeling」ボタン612を押下すればよい。

【0071】指定したラベルが貼り付けられている情報リソースの一覧を参照したい場合、例えば、「料理」というラベルの貼り付けられている情報リソースの一覧を参照したいときには、パレット610の画面上でラベル「料理」を選択し「Finder」というボタン613を押下すればよい。すると、後述するファインダ620が起動されて、ラベル「料理」が貼り付けられている情報リソースの（IDの）一覧がファインダ620によって呈示される。

【0072】複数のラベルを選択した場合は、選択した全てのラベルが貼り付けられている情報リソースの一覧が呈示される。第1の実施形態で説明したように、ラベルに重要度を設定する場合、パレット610の画面上で行うことが望ましい。例えば、パレット610の画面上所望のラベルを選択し、当該ラベルのプロパティ設定画面を開く。そして、例えば5段階の重要度（重要度が最も高い場合を「5」、最も低い場合を「1」として、各ラベル毎に当該プロパティ設定画面上で設定すればよい。

【0073】また、ラベルのプロパティとしては、上記重要度の他に、当該ラベルを付した情報リソースへのアクセス可能な範囲を限定するためのアクセス許可情報があってもよい。例えば、プロパティとして「アクセス可」と設定されているラベルの付された情報リソースに対しては、検索対象とすることができ、「アクセス不可」と設定されているラベルの付された情報リソースに対しては検索対象とすることができない、といった操作も可能とある。さらに、ラベルのプロパティとして、当該ラベルの付された情報リソースに対し、アクセス可能な（当該情報リソースを検索対象とできる）ラベリング集合を限定したり、逆に、アクセスができないラベリング集合を指定したりすることもできよう。これらラベルのプロパティ設定は、上記同様、パレット610の画面上所望のラベルを選択し、当該ラベルのプロパティ設定画面を開いて、当該画面上で行うことが望ましい。

【0074】次に、図9のコレクタ作成処理部504に

ついて説明する。図14において、あるユーザがもっているラベリング集合の中で、たとえば「香港映画」と「写真」というラベルがともに付いた情報リソースをしばしば参照したり、関連した情報を検索したい場合には、対応するラベルの論理式によって特徴づけられる仮想的なフォルダを作成し、それを開けばそのラベルの論理式にマッチする情報リソースのリストが得られるようにしておく。この仮想的なフォルダをコレクタと呼んでいる。またそのときのコレクタを特徴づけるラベルの論理式をコレクタ検索式と呼ぶ。

【0075】図9のコレクタ検索式設定部53では、このコレクタ検索式を設定するものである。図14において、コレクタ710は、「香港映画 and 写真」というコレクタ検索式を持ち、このコレクタ710を開くと情報リソース{a, c}が入っていることがわかる。情報検索時には、このコレクタごとに検索対象のラベリング集合（例えば、「人」）の指定や、検索のためのパラメータ設定を行う。

【0076】図15に示すように、あるコレクタ830の中身をファインダ620で表示したときには、コレクタの中身は2つに分れている。1つはメインリスト810であり、もう1つは候補リスト820である。

【0077】メインリスト810とは、検索発信源であるラベリング集合中の情報リソースのうち、当該コレクタに属する情報リソースのリストであり、候補リスト820とは、コレクタに属する候補となる他のラベリング集合から検索された情報リソースのリストである。他のラベリング集合からの情報検索結果はすべてこの候補リスト820に入り、ユーザの判断で必要に応じてメインリストに移動することができる。

【0078】なお、コレクタにも名前をつけることができる。図14のコレクタ710には、「香港の映画」という名前が付けられており、図15のコレクタ631には、「健康的な料理」という名前が付けられている。

【0079】図9のメインリスト取得部540では、コレクタ検索式設定部530で設定されたコレクタ検索式からメインリスト810を作成する。メインリスト操作部550では、メインリスト810から特定の情報リソースを取り除いたり、情報検索の結果が入った候補リスト820から特定の情報リソースをメインリストに移動させたりする操作を行うようになっている。

【0080】コレクタ作成処理部504で作成されたコレクタは、データ格納部506に格納される。図12は、コレクタを作成する際に、ユーザインタフェース部501に呈示されるファインダ620の基本画面の一例を示したものである。例えば、図12のメニューバー625にある「Collector」をマウスでクリックすることにより、まず、コレクタのフォルダと当該コレクタのアイコンとが新規に作成されて基本画面上に呈示される。その呈示されたアイコンをダブルクリックする

とファインダ620が起動されて、図12に示したような上下2つに分かれた画面が呈示される。上がメインリストの呈示領域623で、下が候補リストの呈示領域624である。コレクタ検索式設定領域621にコレクタ検索式をラベルの論理式で入力すると、コレクタ検索式設定部530は、それをメインリスト取得部540へ渡す。メインリスト取得部540では、当該コレクタ検索式に当てはまる情報リソースをローカルなラベリング集合（例えば、当該ユーザにてラベリングを行ったラベリング集合で、データ格納部506に格納されている）から検索し、その結果得られた情報リソースの集合を羅列したメインリストを作成し、それをメインリスト呈示領域623に呈示する。

【0081】メインリスト呈示部623には、情報リソースのタイトルと、当該情報リソースのID（例えばURL）と、当該情報リソースに付けられているラベルとがメインリストとして呈示されている。

【0082】ユーザがメインリスト呈示領域623に呈示されたメインリストに対し為した操作に基づき、メインリスト操作部550は情報リソースの追加・削除を行い、ユーザの観点に合うようメインリスト中の情報リソースの集合を変更することができる。

【0083】第1の実施形態で説明したように、情報リソースに重要度を設定する場合、ファインダ620の画面上で行うことが望ましい。例えば、図18に示すように、メインリスト呈示領域623に各情報リソースの重要度を設定するための重要度設定領域626を設ける。例えば5段階の重要度（重要度が最も高い場合を「5」、最も低い場合を「1」）として、各情報リソース毎に重要度設定領域626に所望の数値を設定すればよい。

【0084】また、メインリストの上位に表示される情報リソースほど、重要度が高いと予め設定されていてもよい。図12では、「○○○のページ」というタイトルの情報リソースの重要度が最も高く、「健康づくりのページ」というタイトルの情報リソースの重要度が最も低い。この場合、例えばコレクタとしての情報リソースの利用頻度の統計を随時とっておき、利用頻度の高い情報リソースには高い重要度を自動的に設定するようにしてもよい。コレクタを呈示する際には、利用頻度の高いもののほどメインリストの上位に呈示する。また、ユーザがメインリスト呈示領域623に呈示されたメインリストを操作して、メインリストでの情報リソースの呈示順序を変更することにより、情報リソースの重要度を設定するようにしてもよい。このような情報リソースに対する重要度の設定処理は、例えばメインリスト操作部550にて行うようにしてもよい。

【0085】次に、図9の情報検索処理部505について説明する。情報検索処理部505では、図12に示したようなコレクタの選択された状態、すなわち、コレク

タの検索式の設定と、そのコレクタの検索式に基づき作成されたメインリストがメインリスト呈示領域623に呈示されている状態から、例えば、メニューバー621中の「Search」をマウスでダブルクリックされると、送受信部507を介してネットワーク400上のラベリングサーバ411へ、コレクタ（少なくとも、メインリスト）を渡し、検索要求を行う。また、当該情報検索処理部505の情報検索部570において、当該コレクタを用いて検索を実行する。

【0086】なお、検索要求を行う際に、図14に示したように、コレクタに検索範囲を設定してもよい。ラベリングサーバ411は、例えば、図17に示すような構成である。なお、図1と同一部分には同一符号を付している。ラベリングサーバ411では、ネットワーク400、送受信部60を介して、検索用情報リソースリスト取得部10でコレクタを受け取り、ラベリング集合選択部20で、当該コレクタに検索範囲が設定されているときに、検索の対象となるラベリング集合を選択する。情報検索部30において、図3に示すようにラベルスコアの算出処理、情報リソーススコアの算出処理、統合スコアの算出等を行って、候補情報リソースのリスト（候補リスト）を作成し、検索要求元の情報検索装置に送受信部60、ネットワーク400を介して送り返す。情報検索装置では、図13に示すように、当該候補リストを候補リスト呈示領域624に呈示する。

【0087】このように、コレクタに予めコレクタ検索式を設定し、ファインダ620から検索を実行することで、ラベリングサーバ411からコレクタ検索式にマッチする情報リソースを取ってくるできるようになっている。取ってきた情報リソースは、候補リストとして別個呈示されるとともに、コレクタに格納されるので、ユーザがカスタマイズしたメインリストを乱すことはない。ユーザは、候補リストとして候補リスト呈示領域624に呈示された情報リソースのうち必要なものだけをメインリストに移し替えばよい。

【0088】コレクタ選択部560では、データ格納部506に格納されたあるいはコレクタ作成処理部504で作成されたコレクタのうち、検索のためのコレクタを選択するためのものである。

【0089】情報検索部670は、例えば、図1の情報検索処理部1と同様な構成であって、検索用情報リソースリスト取得部10でコレクタ選択部560で選択されたコレクタを受け取って、第1の実施形態で説明したようにして、データ格納部506から情報リソースの検索を行う（図1の検索用情報リソースリスト取得部10、ラベリング集合選択部20、情報検索部30の処理動作と同様）。

【0090】候補情報リソースリスト呈示部580では、情報検索部570で検索した結果得られた候補情報リソースのリストを呈示するためのものである。次に、

図16を参照して、ラベリング集合Aから作成されたコレクタ907を用いて別の2つのラベリング集合B、Cに対して情報検索を行う場合を例にとり、候補情報リソースのスコア計算について具体的に説明する。ユーザA、B、Cは、それぞれの観点からラベルつけた結果、ラベリング集合A、ラベリング集合B、ラベリング集合Cといった分類が得られたとする。

【0091】例えば、ラベリング集合Aには、リソースIDがdat-03の情報リソースにはラベルA1とラベルA3とが付され、リソースIDがdat-10の情報リソースにはラベルA1とラベルA2とが付されている。以下、同様にして、図16のラベリング集合A~Bの表記において、情報リソースに付されているラベルを示すために、リソースIDとラベルの交わる位置に○印を記している。

【0092】このとき、ユーザAは、コレクタ907を生成して、コレクタ検索式として「A1 and A2」を与えたとき、自分自身のラベリング集合内でこのコレクタ検索式にマッチする情報リソースの集合は、{dat-10, dat-36, dat-42}である。すなわち、コレクタ907のメインリスト909には、{dat-10, dat-36, dat-42}が書き込まれている。

【0093】図16では、このコレクタ検索式にマッチする情報リソースのリソースIDとそれに付されたラベルとの交わる位置に記された○印を黒く塗りつぶして示している。

【0094】ラベリング集合Bにおいて、ヒットした（すなわち、コレクタ907にて指定された情報リソースのうちのいずれかに一致する）情報リソースは、dat-10、dat-36で、dat-10はラベルB1とB3とが付されて、dat-36はラベルB3が付されている。これら各ラベルについて加算する数値を全て「1」とすると、ラベルB1のラベルスコアはdat-10のみがヒットしているので「1」、ラベルB3のラベルスコアはdat-10とdat-36とがヒットしているので「2」となる。

【0095】一方、ラベリング集合Cにおいて、ヒットした（すなわち、コレクタ907にて指定された情報リソースのうちのいずれかに一致する）情報リソースは、dat-10、dat-36、dat-42で、dat-10はラベルC2とC3とが付され、dat-36はラベルB2、C3、C4が付され、dat-42はラベルC2が付されている。これら各ラベルについて加算する数値を全て「1」とすると、ラベルC2のラベルスコアはdat-10、dat-36、dat-42がヒットしているので「3」、ラベルC3のラベルスコアはdat-10、dat-36がヒットしているので「2」、ラベルC4のラベルスコアはdat-10のみがヒットしているので「1」となる。

【0096】コレクタ907を用いてラベリング集合Bから検索された候補情報リソースの集合は、{dat-20, dat-22, dat-53, dat-60}である。また、コレクタ907を用いてラベリング集合Cから検索された候補情報リソースの集合は、{dat-04, dat-19, dat-22, dat-60}である。図16のラベリング集合A、Bの表記では、各ラベリング集合から検索された候補情報リソースのリソースIDとそれに付されたラベルとの交わる位置に二重マルを記している。

【0097】ラベリング集合Bにおいて、候補情報リソースdat-20、dat-22、dat-53、dat-60のそれぞれの情報リソーススコアを計算する。情報リソースdat-20の情報リソーススコアはラベルB3のラベルスコア「2」となり、情報リソースdat-22の情報リソーススコアはラベルB1のラベルスコア「1」となり、情報リソースdat-53の情報リソーススコアはラベルB1のラベルスコア「1」とラベルB3のラベルスコア「2」とを加算して「3」となり、情報リソースdat-60の情報リソーススコアはラベルB1のラベルスコア「1」となる。すなわち、ラベリング集合Bにおける

dat-20の情報リソーススコアは「2」

dat-22の情報リソーススコアは「1」

dat-53の情報リソーススコアは「3」

dat-60の情報リソーススコアは「1」

である。

【0098】一方、ラベリング集合Bにおいて、候補情報リソースdat-04、dat-19、dat-22、dat-60のそれぞれの情報リソーススコアを計算する。情報リソースdat-04の情報リソーススコアはラベルC4のラベルスコア「1」となり、情報リソースdat-19の情報リソーススコアはラベルC3のラベルスコア「2」とラベルC4のラベルスコア「1」を加算して「3」となり、情報リソースdat-22の情報リソーススコアはラベルC2のラベルスコア「3」とラベルC3のラベルスコア「2」とを加算して「5」となり、情報リソースdat-60の情報リソーススコアはラベルC2のラベルスコア「3」となる。すなわち、ラベリング集合Bにおいて、

dat-04の情報リソーススコアは「1」

dat-19の情報リソーススコアは「3」

dat-22の情報リソーススコアは「5」

dat-60の情報リソーススコアは「3」

である。

【0099】次に、コレクタ907を用いてラベリング集合B、Cから検索された候補情報リソースの統合スコアを求める。ここでは、ラベリング集合BとCの重み係数とともに「1」とする。すると、情報リソースdat-20、dat-22、dat-53、dat-60、

dat-04、dat-19のそれぞれの統合スコアはラベリング集合Bでの情報リソーススコアとラベリング集合Cでの情報リソーススコアとを加算して、情報リソースdat-20の統合スコアは「2」+「0」=「2」
 情報リソースdat-22の統合スコアは「1」+「5」=「6」
 情報リソースdat-53の統合スコアは「3」+「0」=「3」
 情報リソースdat-60の統合スコアは「1」+「4」=「5」
 情報リソースdat-04の統合スコアは「0」+「1」=「1」
 情報リソースdat-19の統合スコアは「0」+「3」=「3」

となる。統合スコアの値の大きいものから順に降順に並べて替えて、図16に示すような候補情報リソースを順位付けした候補リスト910が得られる。候補リスト910は、コレクタ907に保存される(図15参照)。

【0100】以上説明したように、上記第2の実施形態によれば、ネットワーク上の複数のユーザが参照可能な文書や画像といった任意の情報リソースに対して、各ユーザが自由な観点に基づいて収集し、ラベル付け(ラベリング)した結果得られる分類情報を利用することにより、関心のある情報と関連する未知の情報リソースをネットワークを介して効率良く検索および収集することができる。

【0101】なお、ネットワークを介した情報検索に限らず、オフラインでの情報検索もちろん可能である。すなわち、図9の情報検索装置のデータ格納部506には、ユーザの収集した情報リソースと、その情報リソースに対し当該ユーザが独自にラベリングした結果としてのラベルに関する情報と、コレクタに関する情報等が格納されているので、当該ユーザが複数のラベリング集合を持っているのであれば、そのうちの1つのラベリング集合からラベルの論理式にて特徴付けられる1または複数の検索用情報リソースを選択し、これを基に、他のラベリング集合から当該検索用情報リソースに関連する情報リソースを検索して呈示するようにしてもよい。

【0102】

【発明の効果】以上説明したように、本発明によれば、ユーザは主観的なラベリングによる自由度の高い情報整理をしながら、集めた情報と関連する情報を容易に検索することが可能となる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態に係る情報検索方法を用いた情報検索装置の基本的な構成を概略的に示した図。

【図2】2つのラベリング集合間で情報検索を行う様子を概念的に示した図。

【図3】ラベルスコア算出部と情報リソーススコア算出部32の処理動作を示したフローチャート。

【図4】図3のステップS2からステップS4までの処理(ラベルスコアの算出処理)を説明するための概念図。

【図5】図3のステップS5からステップS7までの処理(情報リソーススコアの算出処理)を説明するための概念図。

【図6】複数(例えば4つ)のラベリング集合A~Dから情報リソースを検索する場合を概念的に示したもので、個々の検索対象のラベリング集合で得られた情報検索の結果を統合する場合を説明するための図。

【図7】本発明の情報検索方法を用いて情報を検索および収集する場合を概念的に示した図。

【図8】本発明の第2の実施形態に係る情報検索装置を用いて、例えば、ネットワークを介して情報検索を行う場合を概念的に示した図。

【図9】本発明の第2の実施形態に係る情報検索装置の要部の構成を概略的に示した図。

【図10】図9の情報検索装置の動作を、ユーザインタフェース部にて呈示されるユーザインタフェース画面(ラベルの操作を行うパレットと、ラベルやコレクタごとに関係した情報リソースのリストを表示するファインダ)と、主に情報検索のために利用するコレクタとの関連に基づき概念的に示した図。

【図11】ユーザインタフェース部にて呈示されるパレットの画面呈示例を示した図。

【図12】コレクタを作成する際に、ユーザインタフェース部にて呈示されるファインダの呈示画面の一例を示した図で、検索結果(候補リスト)を呈示していない場合を示している。

【図13】コレクタを作成する際に、ユーザインタフェース部にて呈示されるファインダの呈示画面の一例を示した図で、検索結果(候補リスト)を呈示した場合を示している。

【図14】情報検索の際に用いるコレクタについて説明するための図。

【図15】コレクタの中身をファインダで表示したときの表示例。

【図16】ラベリング集合Aから作成されたコレクタを用いて別の2つのラベリング集合B、Cに対して情報検索を行う場合を例にとり、候補情報リソースのスコア計算について具体的に説明するための図。

【図17】ラベリングサーバの構成を概略的に示した図。

【図18】各情報リソースの重要度を設定するための重要度設定領域を有したファインダの画面呈示例を示した図。

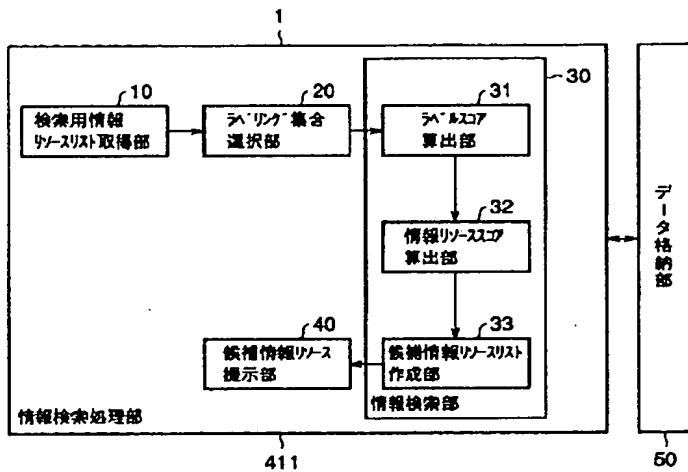
【符号の説明】

1…情報検索処理部

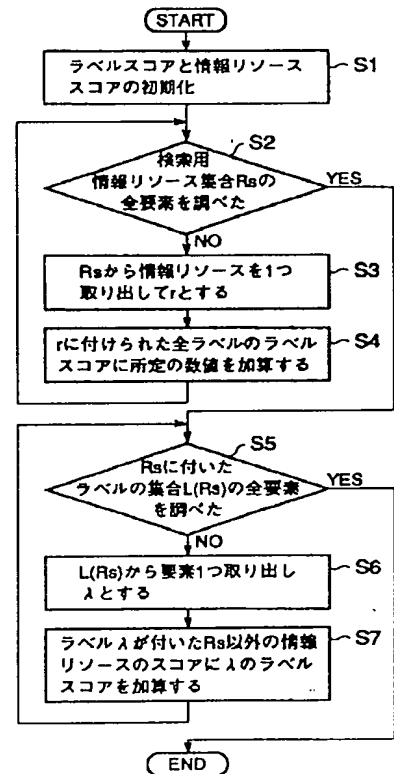
10…検索用情報リソースリスト取得部
 20…ラベリング集合選択部
 30…情報検索部
 31…ラベルスコア算出部
 32…情報リソーススコア算出部
 33…候補情報リソースリスト作成部
 40…候補情報リソース提示部
 411…ラベリングサーバ
 412…ラベリングデータベース
 501…ユーザインタフェース部
 502…処理部
 503…ラベリング処理部
 504…コレクタ作成処理部
 505…情報検索処理部

506…データ格納部
 507…送受信部
 510…ラベル定義部
 520…ラベル付与部
 530…コレクタ検索式
 540…メインリスト取得部
 550…メインリスト操作部
 560…コレクタ選択部
 570…情報検索部
 580…候補情報リソースリスト提示部
 610…パレット
 620…ファインダ
 630、710…コレクタ

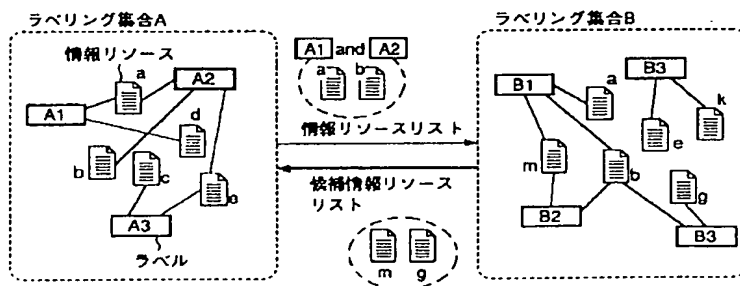
【図1】



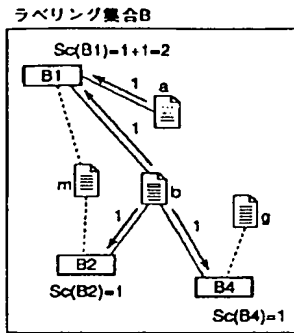
【図3】



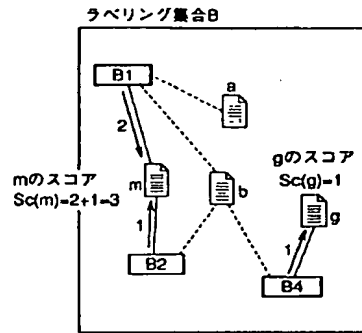
【図2】



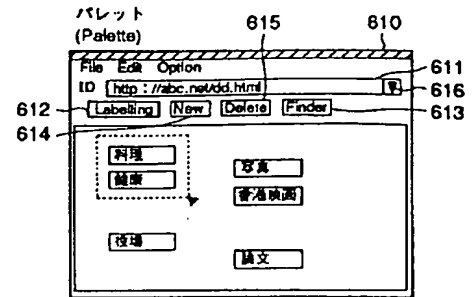
【図4】



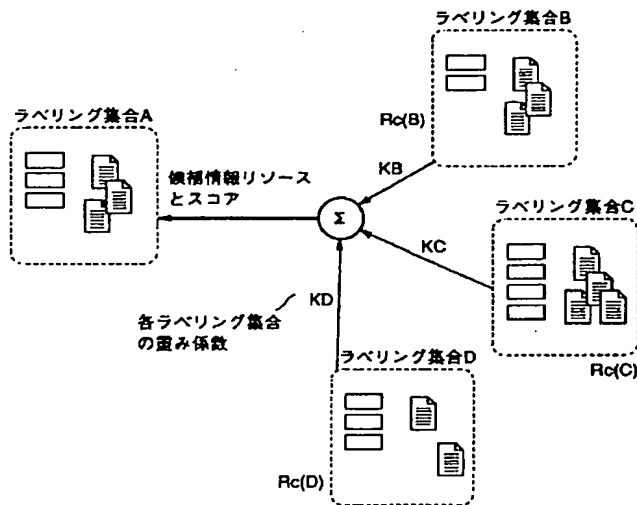
【図5】



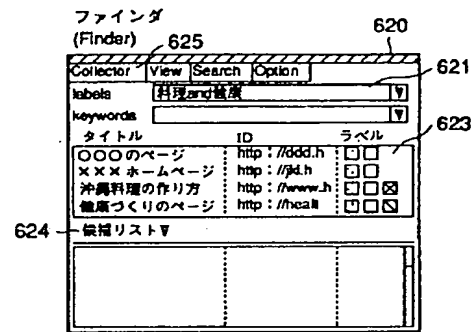
【図11】



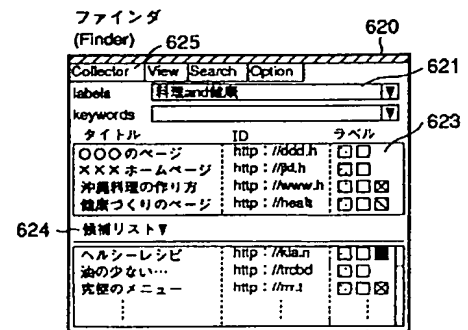
【図6】



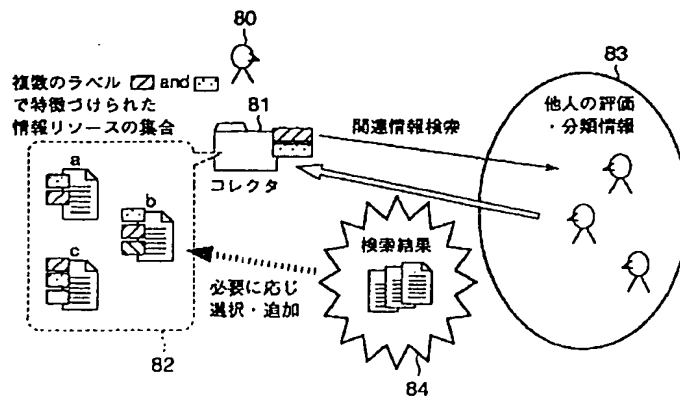
【図12】



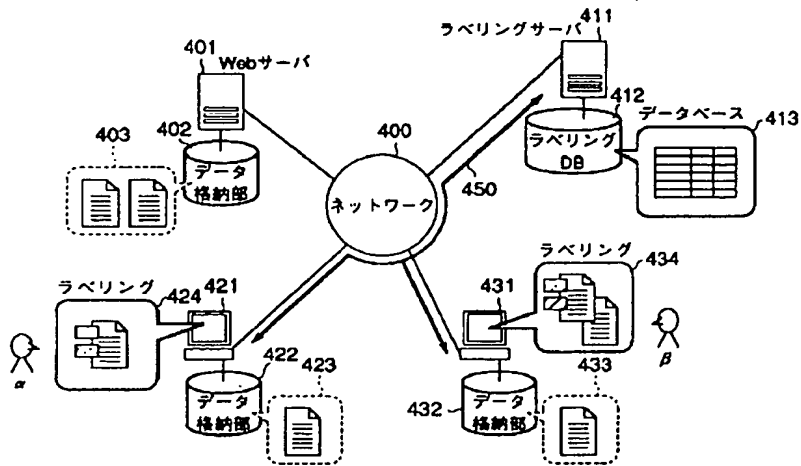
【図13】



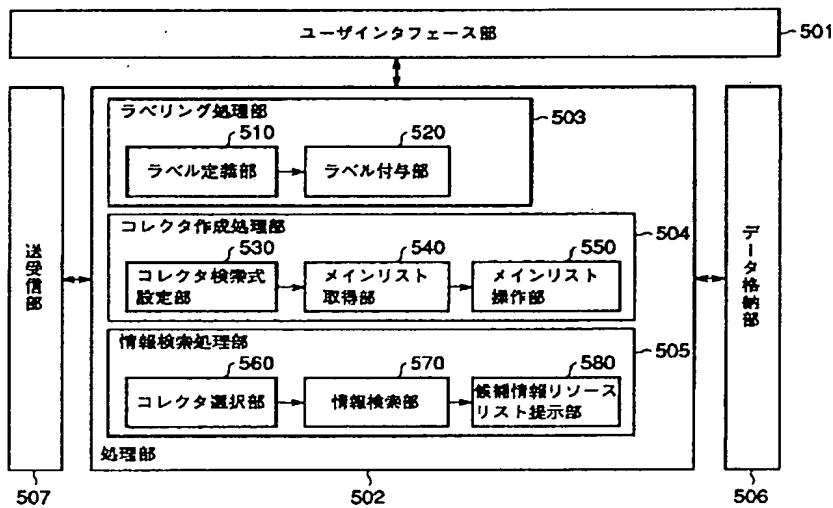
【図7】



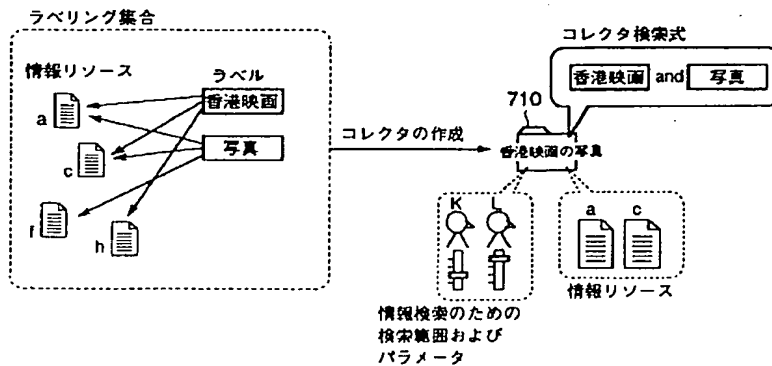
【図 8】



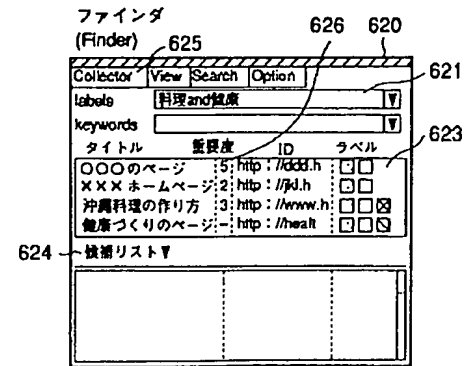
【図 9】



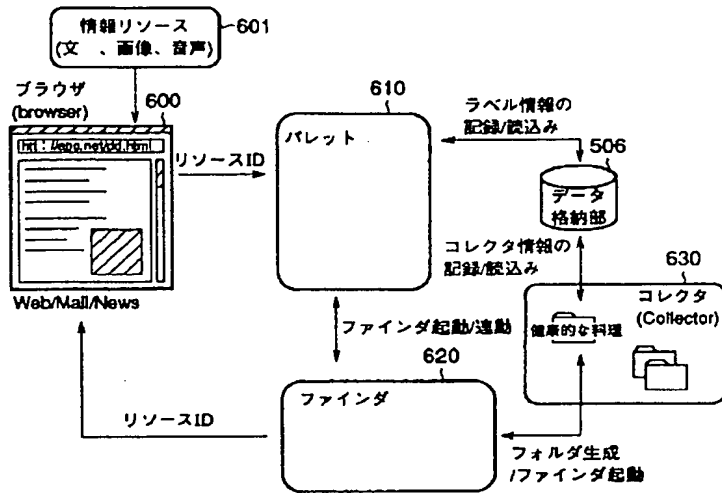
【図 14】



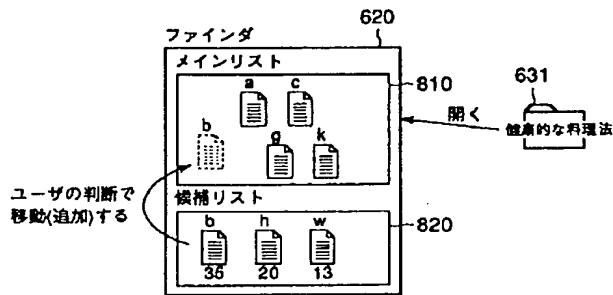
【図 18】



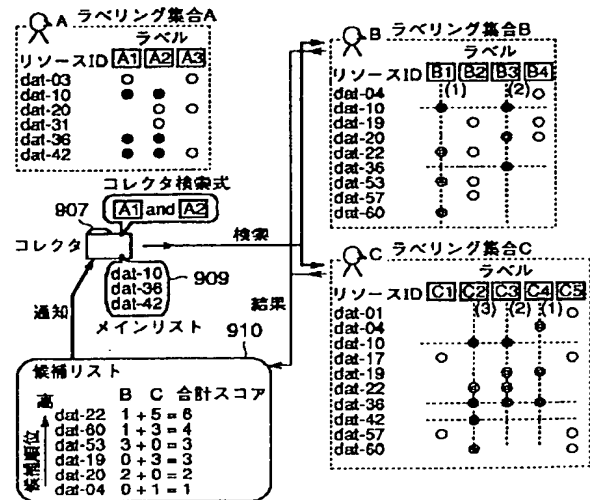
【図10】



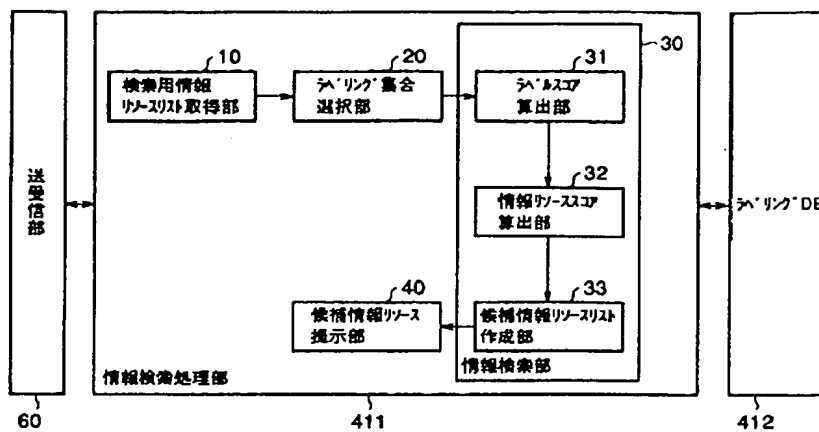
【図15】



【図16】



【図17】



フロントページの続き

(72)発明者 横田 健彦
 神奈川県川崎市幸区柳町70番地 株式会社
 東芝柳町工場内

Fターム(参考) 5B075 ND16 NK04 NK25 PP02 PP03
 PP13 PQ02 PQ46 PR06 QM08